

Model Prediksi Usia Biologis Menggunakan Metode Support Vector Regression dan Principal Component Analysis pada Biomarker Klinis = Biological Age Prediction Model Using Support Vector Regression and Principal Component Analysis Methods on Clinical Biomarkers

Annisa Zahra, author

Deskripsi Lengkap: <https://lib.ui.ac.id/detail?id=9999920539040&lokasi=lokal>

Abstrak

Penuaan adalah proses alami yang secara bertahap menurunkan kondisi fisik dan menyebabkan kemunculan berbagai penyakit, yang pada akhirnya dapat mengurangi rentang hidup makhluk hidup serta berujung pada kematian. Dalam konteks ini, usia biologis berperan sebagai indikator penting yang mampu mengevaluasi proses penuaan dan prediksi penyakit lebih efektif dibandingkan dengan usia kronologis. Hal ini dikarenakan usia biologis juga memperhatikan kondisi fisiologis individu, bukan hanya mengukur lamanya hidup seseorang sejak lahir. Penelitian ini berfokus pada proses penuaan alami yang tidak dipengaruhi oleh penyakit. Dengan demikian, model ini dapat dijadikan alat untuk mengidentifikasi individu yang jalur penuaannya menyimpang dari jalur penuaan yang sehat. Penelitian ini menggunakan metode Support Vector Regression dan Principal Component Analysis untuk memprediksi usia biologis berdasarkan biomarker klinis yang berkontribusi terhadap proses penuaan. Data yang digunakan pada penelitian ini adalah data medis yang berasal dari Kementerian Kesehatan Republik Indonesia. Pada dataset, dilakukan data preprocessing yang meliputi pengubahan tipe data, penghapusan kolom yang tidak digunakan, penyaringan usia partisipan, pembentukan data sintetis, dan pemisahan dataset pria dan wanita. Selanjutnya, dilakukan feature selection, uji multikolinearitas, dan pembentukan model menggunakan metode Support Vector Regression dan Principal Component Analysis. Performa dari model yang dibentuk, dievaluasi menggunakan Root Mean Squared Error dan Coefficient of Determination. Untuk model yang menggunakan metode Support Vector Regression, didapatkan nilai RMSE = 5, 228 dan $r^2 = 0, 807$ pada model pria, serta nilai RMSE = 1, 798 dan $r^2 = 0, 959$ pada model wanita. Sementara itu, model yang menggunakan metode Principal Component Analysis didapatkan nilai RMSE = 6, 835 dan $r^2 = 0, 751$ pada model pria dan nilai RMSE = 5, 35 dan $r^2 = 0, 874$ pada model wanita. Berdasarkan analisis kinerja model yang dilakukan pada penelitian ini, model dengan metode Support Vector Regression lebih unggul dalam memprediksi usia biologis dibandingkan dengan metode Principal Component Analysis.

.....Aging is a natural process that gradually deteriorates physical condition and leads to the emergence of various diseases, ultimately reducing the lifespan of living beings and leading to death. In this context, biological age acts as an important indicator capable of evaluating the aging process and predicting diseases more effectively than chronological age. This is because biological age also considers an individual's physiological condition, not just measuring the length of time of person's life since birth. This research focuses on the natural aging process that is not influenced by disease. Thus, this model can be used as a tool to identify individuals whose aging path deviates from a healthy aging trajectory. This study uses Support Vector Regression and Principal Component Analysis methods to predict biological age based on clinical biomarkers that contribute to the aging process. The data used in this study are medical data from the Ministry of Health of the Republic of Indonesia. In the dataset, data preprocessing is performed, which includes changing data types, removing unused columns, filtering participant ages, forming synthetic data,

and separating datasets for men and women. Next, feature selection, tests of multicollinearity, and model formation using the Support Vector Regression and Principal Component Analysis methods are conducted. The model formed is evaluated using Root Mean Squared Error and Coefficient of Determination. For the model using the Support Vector Regression method, $RMSE=5,228$ and $r^2=0,807$ were obtained for the men model, while an $RMSE=1,798$ and $r^2=0,959$ were obtained for the women model. Conversely, for the model using the Principal Component Analysis method, an $RMSE=6,835$ and $r^2=0,751$ were obtained for the men model, and an $RMSE=5,35$ and $r^2=0,874$ for the women model. Based on the performance analysis conducted in this study, the model using the Support Vector Regression method outperforms the Principal Component Analysis method in predicting biological age.